



Plateforme Génome et Transcriptome Les Nouveautés

Pauline Heuillard

<http://get.genotoul.fr>



@GeT_Genotoul





Qui sommes nous?

- **GeT répartie sur 5 sites, fait partie de Genotoul**
 - Une spécialité technologique et une communauté scientifique par site
 - Des experts en agronomie, environnement, microbiologie et santé
 - Des compétences en biologie, bioinformatique et biostatistique

- **Programme national France Génomique**



- **Labels :**

- IBISA
- Plateforme stratégique INRA



- **Certifications qualité :**

- ISO 9001-2015
- NFX 50 900
- Propel Illumina





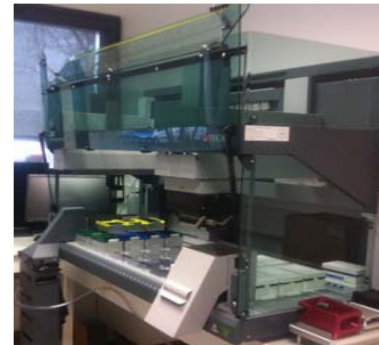
Les Technologies

Les outils périphériques

- **Contrôles qualité des échantillons et des librairies:**



- **Robots pipeteurs pour l'automatisation de la préparation des échantillons:**
 - Partanariat avec Tecan (4 Evo), Agilent Bravo

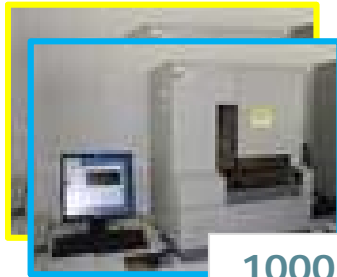


- **Capture de cellules uniques (C1 Fluidigm):**





Les séquenceurs



1000 pb



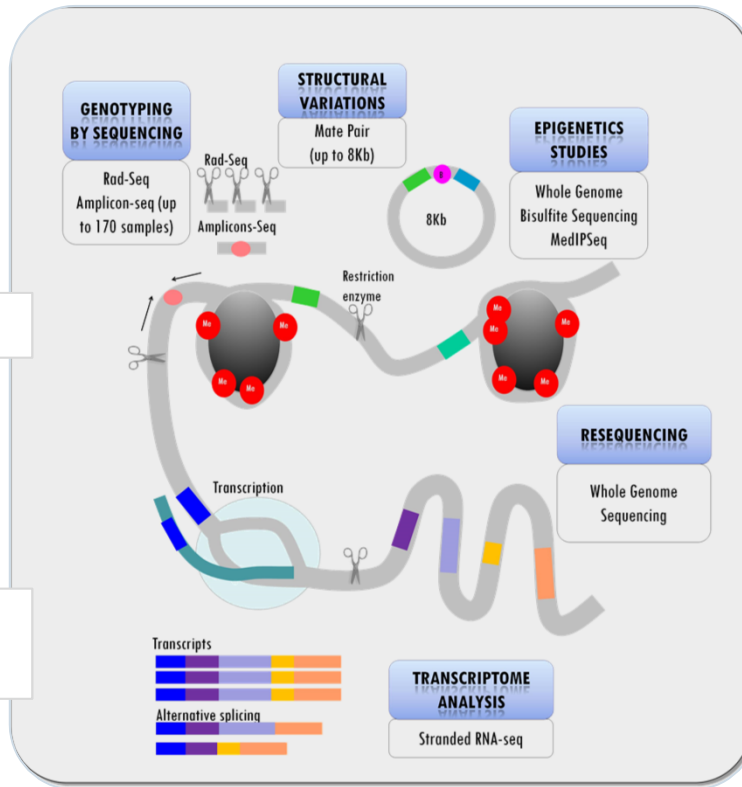
400 pb
1Gb



200 pb
13 Gb



18 000 pb
1,2 Gb (6h)



2x300 pb
15 Gb



2x150 pb
700 Gb





Les Nouveautés

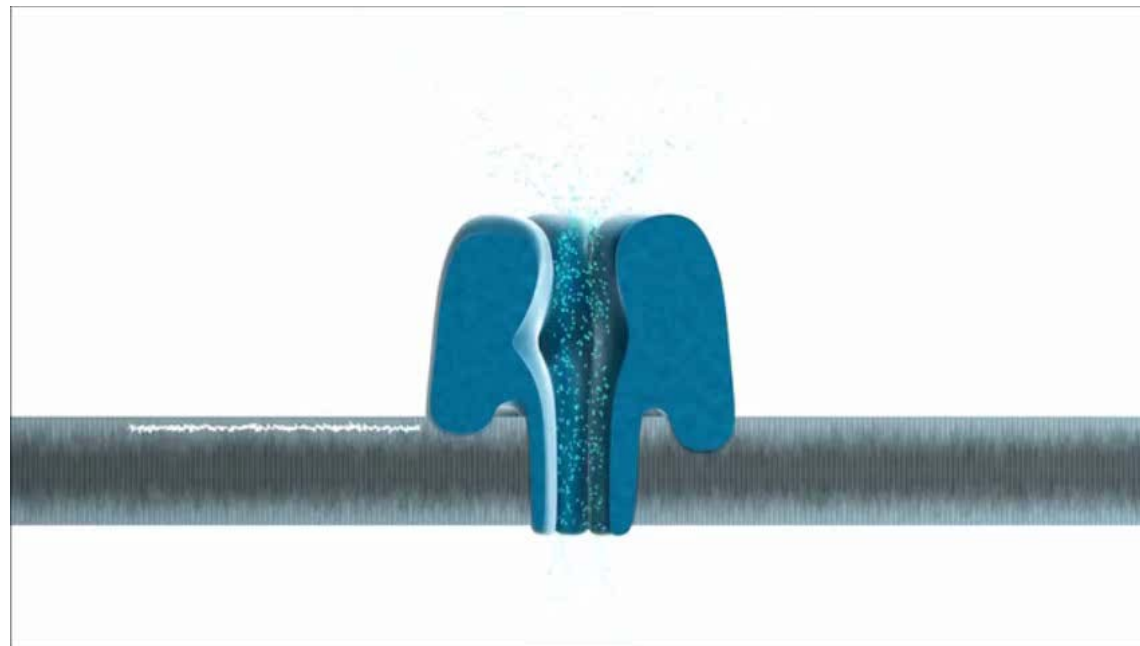


MinION, Oxford Nanopore Technologies

● Technologie de séquençage par électroporation

Potentiel de membrane \Rightarrow Flux ionique et courant électrique mesurable

Encombrement du pore \Rightarrow Variation électrique \Rightarrow Traduction en base

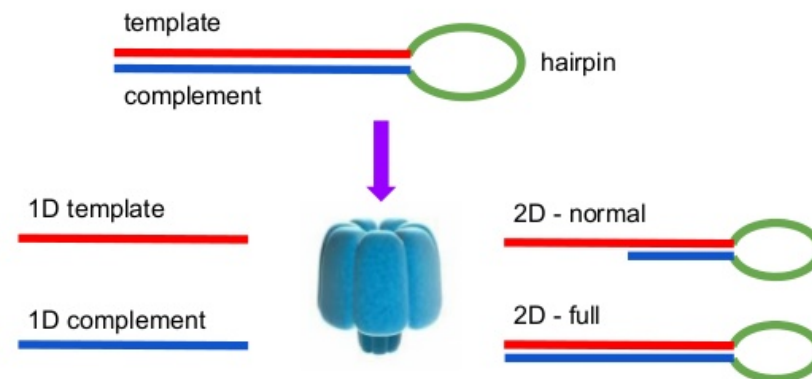




MinION, Oxford Nanopore Technologies

Équipement :

- 3 MinION
- Flowcells de type R9.4 : 512 canaux
- 3 types de kits: rapid, 1D et 2D



- **Input : 1,5 µg d'ADN HPM pour une librairie de ~8 kb**
- **Développement des protocoles en cours grâce à des projets pilotes :**
 - Kit 1D² : augmentation du débit
 - RNA-Seq : Séquençage de l'ARN directement

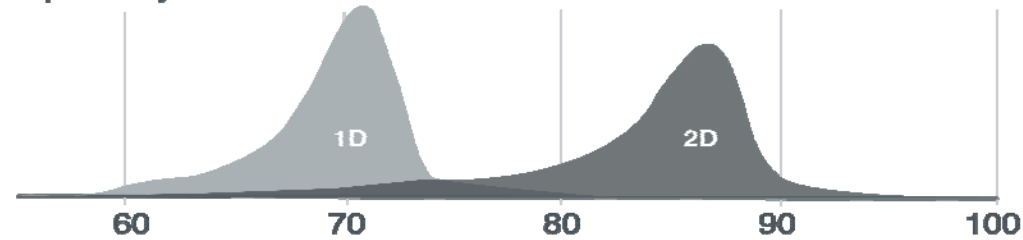
● Basecalling :

- MinKnow
- Nanonet
- Albacore
- Metrichor

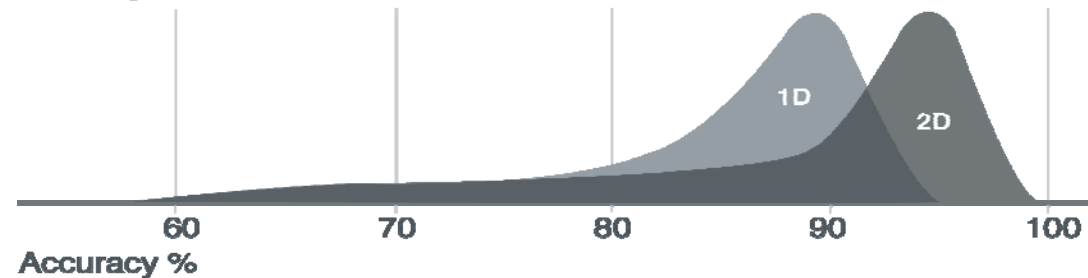


● Taux d'erreur

R7: Up to May 2016



R9: May 2016 Onwards



Données MinION

- 1ères données des runs du projet Bactérie

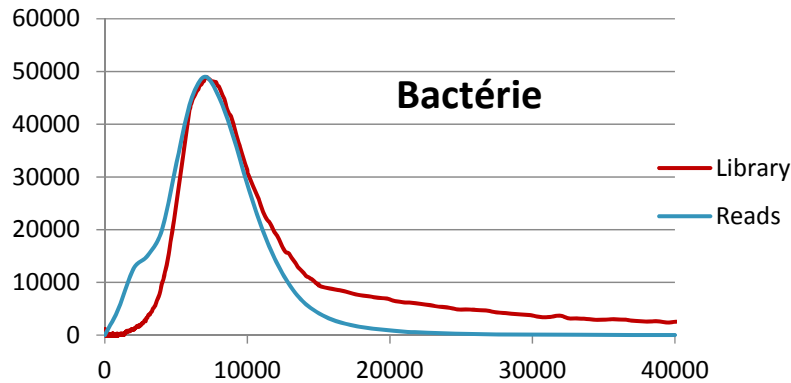
Run	Nb séquences	Qté de données	Taille moyenne	N50	Qualité
Rapid	54 843	368 Mb	6 701	10 326	7,9
1D	355 892	2,6 Gb	7 414	8 427	8,4
2D	97 426	444 Mb	4 558	5 614	15,4
1D proto ++	468 396	7,6 Gb	16 220	18 426	8,3

- Protocoles à adapter...

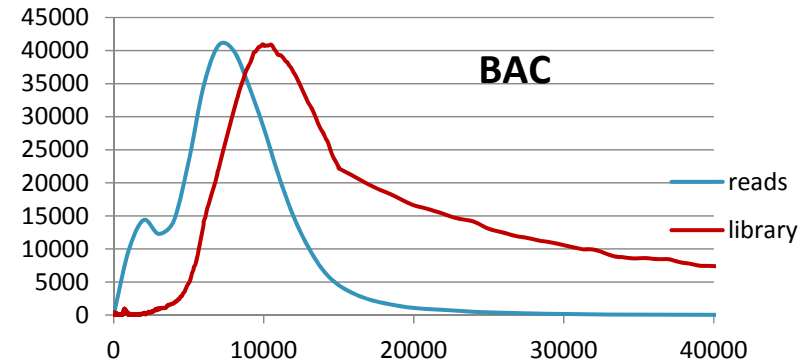


Biais MinION

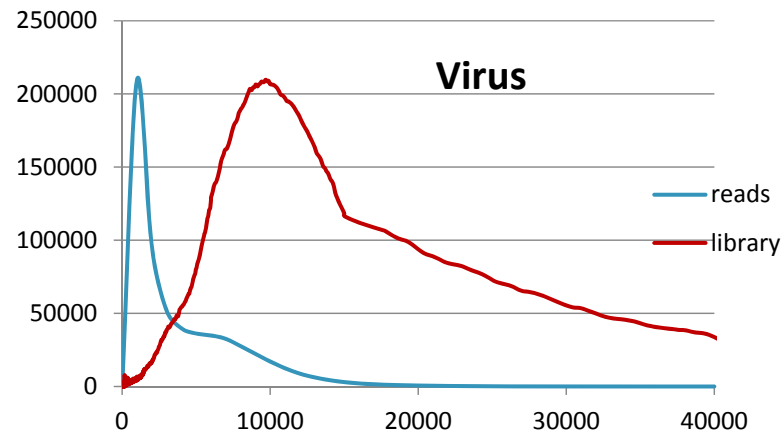
● Impact du profil de la librairie sur la distribution des reads



Librairie homogène => distribution des reads homogène



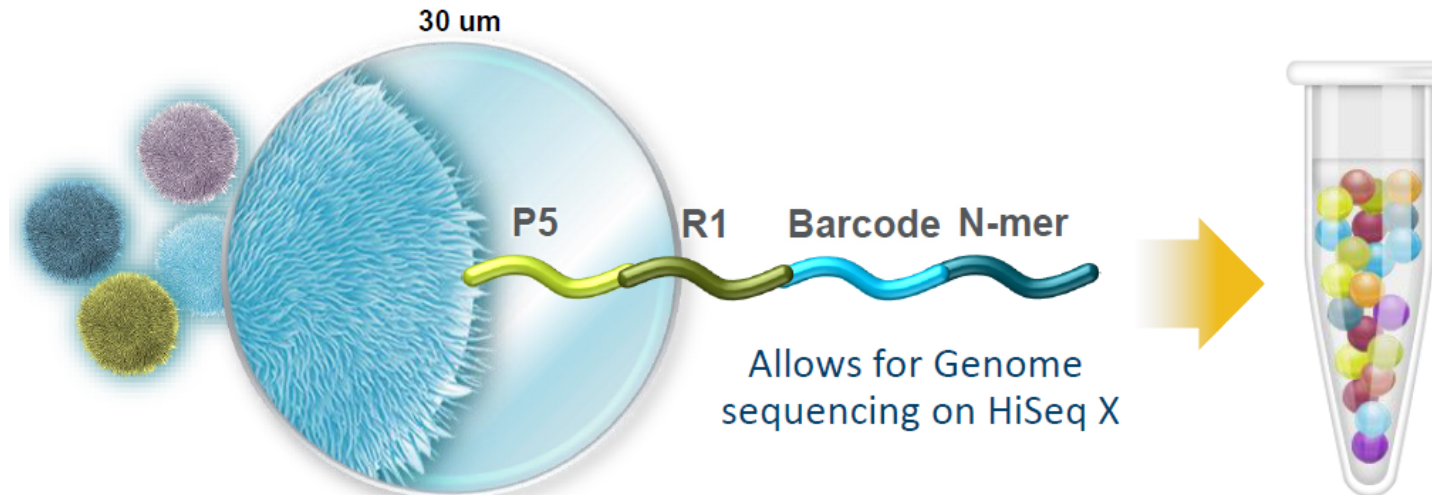
Présence de grands fragments => léger séquençage des petits fragments



Présence de petits et grands fragments => fort séquençage des petits fragments

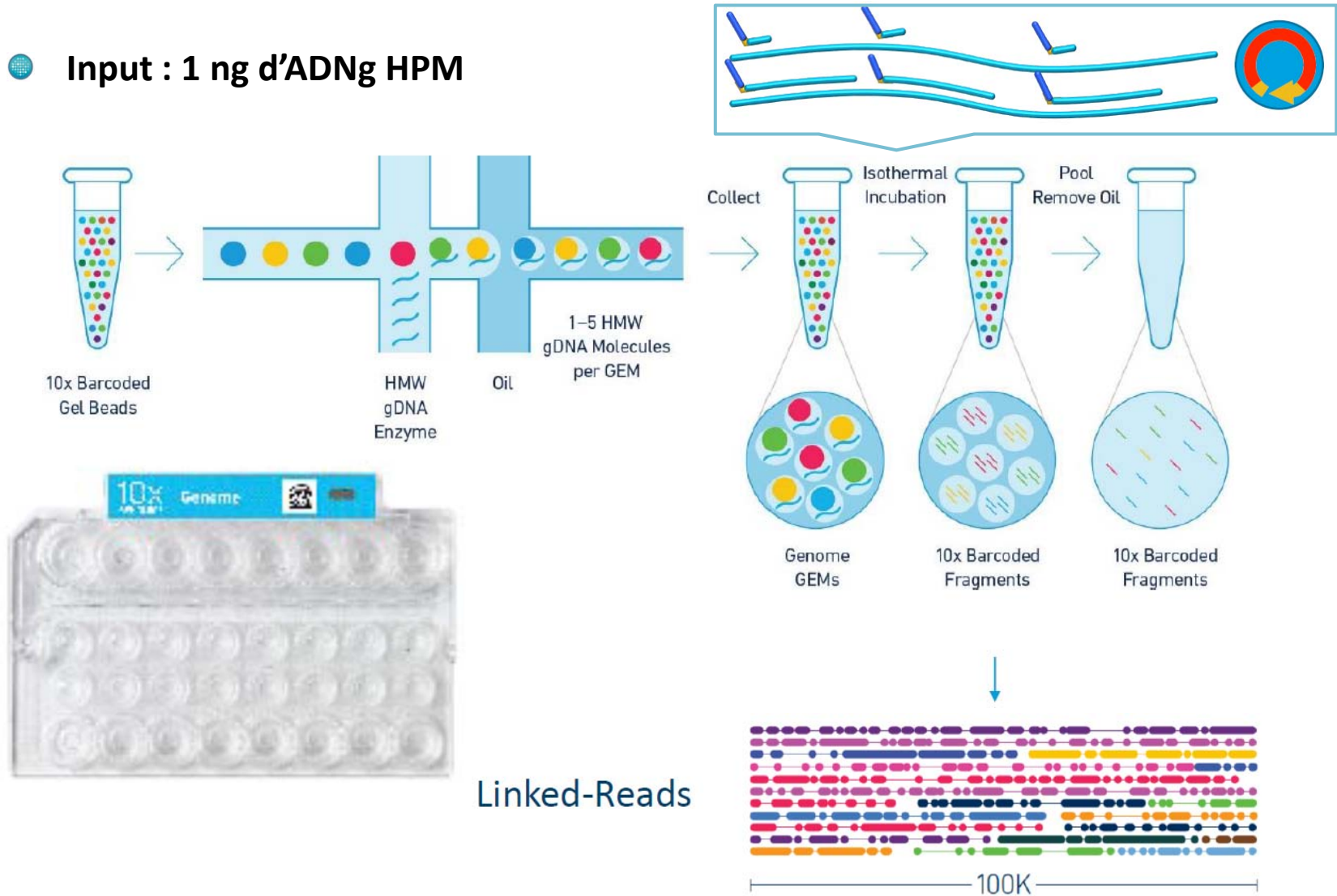
Chromium, 10x Genomics

- Préparation de bibliothèques pour séquençage Illumina
- Génomique à large échelle (>50 kb), haplotypage, détection de variants structuraux, *de novo*
- Analyses de cellules uniques
- Séquençage d'exomes
- 5 sites FG équipés



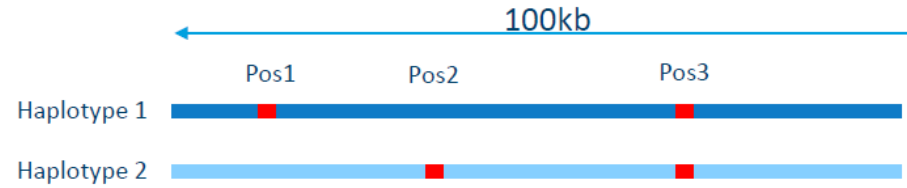
Système 10x Genomics

● Input : 1 ng d'ADNg HPM





Les avantages des linked-reads



Conventional sequencing

Average of two molecules

10x sequencing

Single Molecule Resolution



Les limites de Chromium

- **Technologie développée sur le génome humain (3Gb)**
- **Taille du génome étudié**
 - Minimum 800 Mb
 - Risque d'avoir 2 molécules par GEM qui couvrent le même locus : 1 pour 6000 pour le génome de 3 Gb
- **Taille des fragments**
 - Plus d'une molécule par GEM/barcode
 - Minimum 50 kb
 - 10 molécules aléatoires par GEM
- **Développements à prévoir pour les petits génomes : multiplexage?**



Bioinfo

- **Logiciels:**

Reference-Based
Workflow



Long Ranger
Analysis Pipelines



Loupe
Genome Browser

Reference-**Free**
Workflow



Supernova
De novo Assembly

- **Développement sur génomes diploïdes et ~3 Gb, open-source?**



Genotoul
GeT



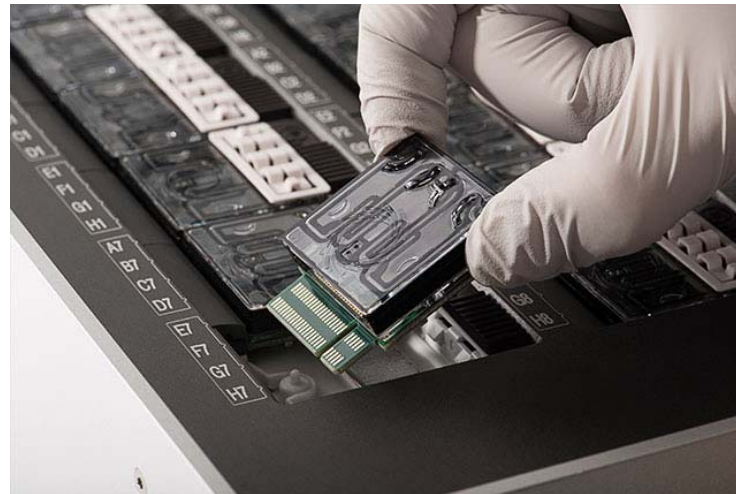
INRA
SCIENCE & IMPACT

À venir

PromethION, ONT

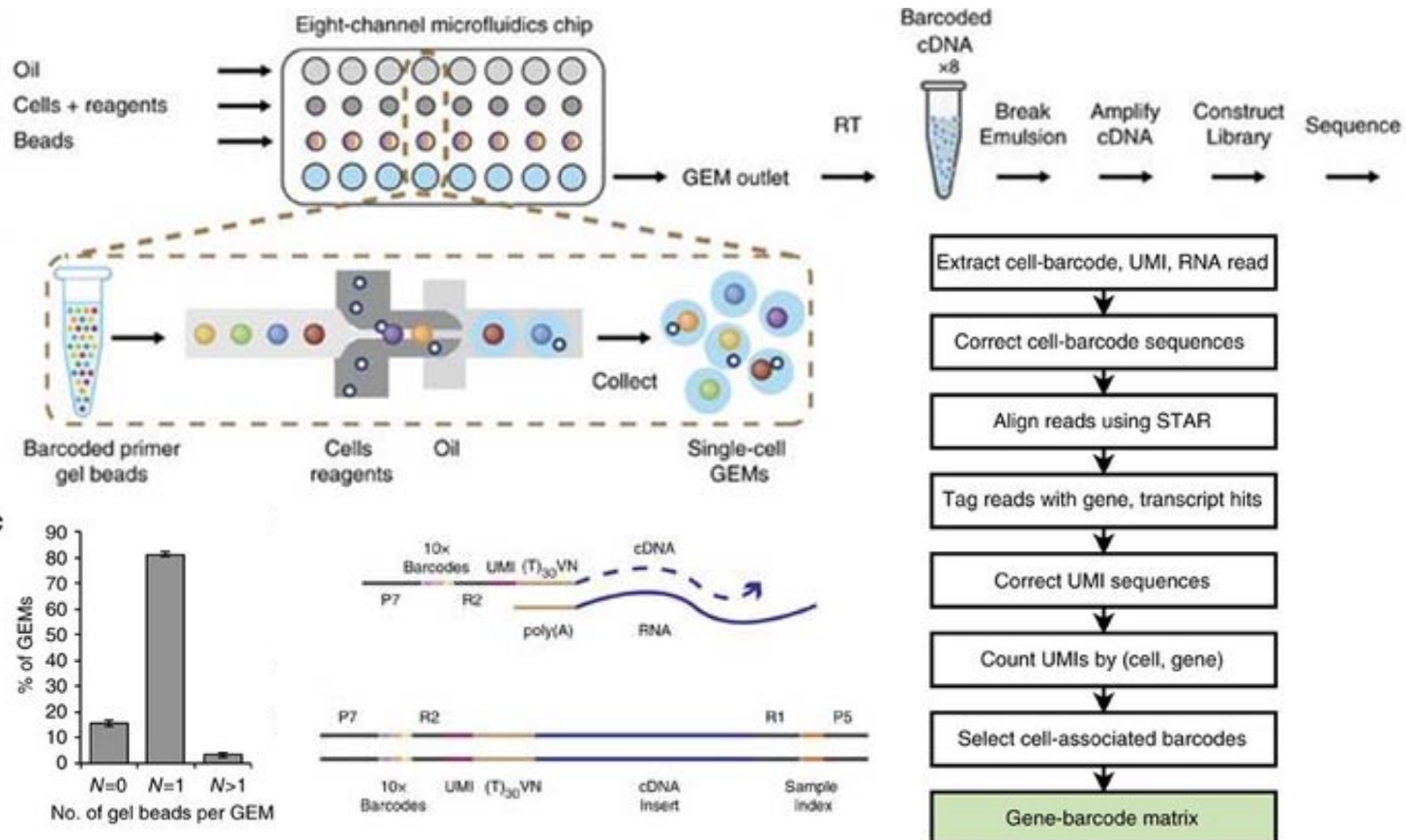
- **Systeme modulable de 1 à 192 échantillons :**

- 48 chips individuelles
- 4 échantillons par chip
- 3000 canaux actifs par chip



Single Cell, 10X Genomics

- 3' RNA-Seq (poly A)
- 700 à 70000 cellules en GEM en 10 min
- Efficacité de capture >65% pour les types cellulaires rares dans les populations hétérogènes,



NovaSeq, Illumina



- 2 plateformes : NovaSeq5000 et NovaSeq6000
- 4 types de FC: S1, S2, S3 et S4
- 100 transcriptomes /run avec S2



Run times:
<1 to ~2.5d
based on
system, FC
and read
length



Configure
output to
match your
application
and study
size

Single flow cell output (1 or 2 can run simultaneously)

Flow Cell Type	NovaSeq 5000	NovaSeq 6000	Reads per Flow Cell	Output (Gb) per Flow Cell		
				100 cycles	200 cycles	300 cycles
S4*		✓	10B			3000
S3*		✓	6.6B			2000
S2	✓	✓	3.3 B	333	666	1000
S1*	✓	✓	1.6 B	167	333	500



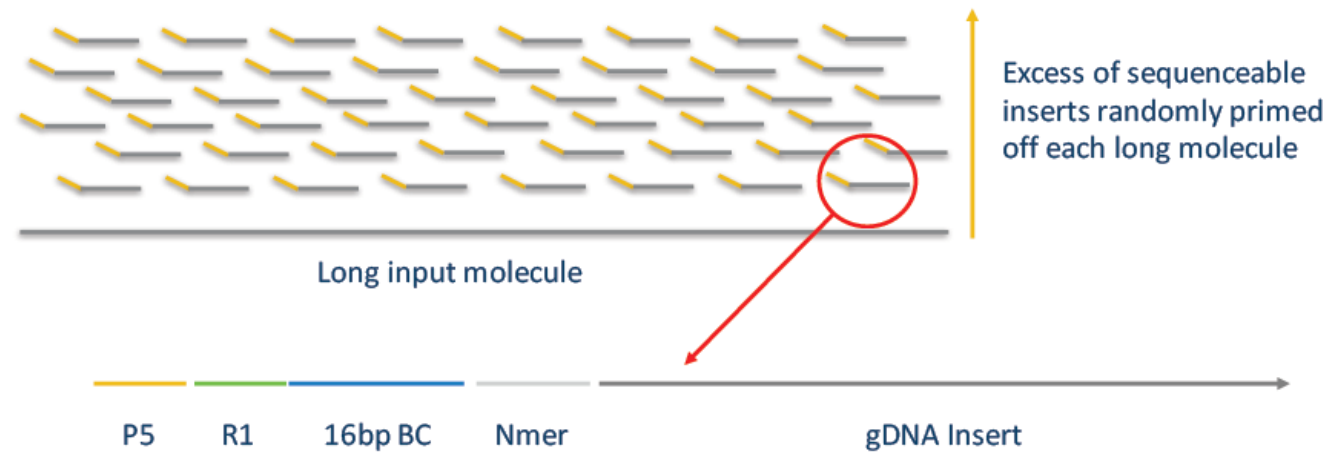
Merci à la plateforme bioinfo



Merci pour votre
attention

Annexe 1

Example – Barcoding a Short Read Library



At 30X read coverage, ~35 library fragments will end up sequenced from each 50Kb input molecule

$35 \times 2 \times 150\text{bp} \approx 10\text{Kb}$, or 0.2X read coverage per molecule

Reads from the same input molecule are called “Linked-Reads”